

al, Susan, 1979. *Language shift: social determinants of change in bilingual Austria*, New York: Academic Press.

bov, William, 1963. The social motivations of a second change. *Word* 19:273-309.

—, 1965. On the mechanism of linguistic change, *Georgetown University Monograph Series on Language and Linguistics*, 18: 91-114, Washington D.C.: Georgetown University Press.

nkoff, G. 1972. Language use in multilingual societies: some alternate approaches. In Pride and Holmes (eds.) *Sociolinguistics*, Harmondsworth: Penguin.

inreich, U., W Labov, and M. Herzog, 1968. Empirical foundations for a theory of language change. In Lehmann and Malkiel (eds.) *Directions for Historical Linguistics*, Austin: University of Texas Press.

## How Native-like? measuring language proficiency in bilinguals

DAVID E. INGRAM  
Darwin Community College, Australia

**ABSTRACT.** The article is in applied bilingualism which raises certain issues regarding the adequacy of test instruments such as the set of Absolute Language Proficiency Ratings of the Foreign Services Institute, School of Language Studies in the U.S.A. specially in the context of new and fast changing socio-political situation (which creates bilingual situations most of the time) in various countries. A direct proficiency rating scale, the ASLPR is developed to have great discrimination, wider applicability and a clearer focus on all four macro skills than the FSI scale provides. The paper outlines the development of the scale, its associated test procedures, the results of the various trials in some countries and the scope and the limitations of the instruments. It is suggested that the ASLPR has a high level of validity and reliability and that it can be confidently used wherever a meaningful statement of a bilingual person's practical skills in the second language is required.

### 1. P R E A M B L E

As second and foreign language teaching have come to focus more on communication goal and communicative methods, the need for new test instruments focussing on communication skills or practical language proficiency has increased. Traditional approaches to language testing were indirect, i.e., they assumed that a measure of one thing (viz., the learner's knowledge or linguistic competence) could serve as a measure of something else (viz., his language proficiency). This, for example, discrete-point tests assess knowledge of individual items, especially, of grammar. A typical item might be as follows:

Tick the item that best fills the gap.

Question: "How old is Bill?"

"He's two years older ----- his sister Mary."

Answer: that ☐  
as ☐  
than ☐  
of ☐

Thus this item tests to see if the learner knows that *than* is the conjunction used when comparisons are made. Semi-direct tests (such as cloze, dictation or white noise) seek to be as objective as indirect tests but incorporate each item in something approximating the total language activity. They are generally interpreted comparatively either by ranking learners or by correlating the results with standardized indirect tests, i.e., with knowledge-based tests.

However, knowledge is not the same as proficiency; a learner can have much knowledge about a language and yet not be able to mobilize that knowledge readily or effectively to carry out communication tasks. Thus, for example, a learner may score well on the TOEFL or CELT tests and yet remain unable to carry out everyday communication tasks or to use the language for academic study (cf. Angelis 1979; Oller and Spolsky 1979). Direct tests, which seek to relate proficiency statements to actual language behaviour, would seem, therefore, to be more informative. Rather than just measure knowledge about the language, such instruments describe different levels of proficiency in terms of the communication tasks or activities that can be carried out and how they are carried out. Learners are rated by being matched against the level on the scale that best describes their language behaviour. In essence, the path of second language development from zero to native-like is envisaged, different points along that path described, and the learner's position along the path identified as his second language behaviour is observed.

The principal instrument of this type that has been available is the "Absolute Language Proficiency Ratings" of the United States Foreign Service Institute School of Language Studies (1968). However, the FSI Scale has a number of limitations: it is primarily intended for use with well-educated civil and foreign service personnel learning to use the language in professional contexts (see Sollenberger 1978; Clark 1972: 121); it contains only two sub-scales (for Speaking, which includes listening, and for Reading); and its six steps from zero to five provide insufficient differentiation at the lower end where most students in classes are found. Thus, the FSI Scale has serious limitations when used in a large-scale migrant education programme catering for learners differing widely in educational background and interested in acquiring, at first,

minimum survival skills. Its limitations are no less severe when it is applied to foreign language learners in the school system since they tend to cluster in the lower proficiency levels and are interested in developing all four macroskills. - *which are what?*

2. DEVELOPMENT OF THE ASLPR

In September 1978, the Australian Department of Immigration and Ethnic Affairs established a national committee, the Joint Commonwealth-States Committee on the Adult Migrant Education Program, to prepare new programmes for use in its large Adult Migrant Education Program. The Joint Committee rapidly recognized the need for a performance-based scale for use in stating the proficiency of learners in its programme and for use as a framework of proficiency development within which a variety of language programmes could be coherently planned. Since the FSI Scale was considered inadequate for these purposes, the Joint Committee initiated research into the development of a new scale. This work, undertaken by Elaine Wylie and the present author, has led to the development of the Australian Second Language Proficiency Ratings (ASLPR). The following developmental procedures were adopted:

1. The FSI Scale was taken as a starting-point since it was the most widely accepted and thoroughly researched performance-based scale then available (see, for example, Clark 1978; Clifford 1978; Graham 1978; Ingram 1977; Lado 1978; Lovelace 1978; Mullen 1978; Sollenberger 1978; Wilds 1965). Despite this starting-point and some surface similarities remaining in some definitions, the FSI Scale and the ASLPR now differ substantially in ways that will emerge later in this paper. From this starting-point and drawing on their experience of second and foreign language teaching totalling some forty years in different teaching situations in a dozen different countries, the researchers described language behaviour at nine proficiency levels from zero to native-like. In addition, they assessed their emerging description of the path of language proficiency development against the findings in the literature of developmental psycholinguistics (summarized, for example in Ingram 1978, Chapter 4).
2. The emerging scale was then re-assessed in some fifty interviews with adult learners of English as a Second Language (ESL) distributed throughout the proficiency range from zero to native-like. These interviews were planned to elicit the characteristics of language behaviour at each level so as to ensure the appropriateness and coherence of the behavioural descriptions adopted. This procedure has since been replicated informally many times over and formally for a further forty adolescent ESL learners, fifty FL learners of French, twenty Italian, and thirty Japanese.
3. The scale was then released for use throughout the Adult Migrant Education Program in Australia so that further feedback on the scale could be provided as a result of its practical implementation.



4. Formal trialling of the ASLPR was then undertaken to determine whether the definitions could reliably and validly be applied to adult ESL learners, to adolescent ESL learners, and to adolescent and adult FL learners of French, Italian and Japanese. These formal trials, which have allowed the appropriateness and coherence of the scale's descriptions to be further assessed, will be reported later in this paper.

### 3. DESCRIPTION OF THE ASLPR

The ASLPR (see Table 3, pp. 58-59) describes language performance at nine points along the continuum from zero to native-like proficiency in each of the four macroskills. An additional three points are available at 2+, 3+ and 4+ for learners, whose proficiency is significantly above one level but not at the next. The outcome of using the ASLPR is a profile (e.g., S:1, L:1+, R:1, W:1-) indicating the learner's proficiency in each macroskill.

The descriptions seek to provide an overall picture of language behaviour at each level of proficiency: it is the rater's global assessment that is important in applying the scale with the learner's proficiency being assigned to the level whose description best matches his observed behaviour. In this, the ASLPR differs fundamentally from the FSI Scale where each definition is a checklist, all the requirements of which must be satisfied before a rating is assigned. Although the ASLPR approach increases the subjectivity of the rating process, it better recognizes that language is highly complex and may develop at slightly different rates in different aspects. Thus, for instance, a Chinese businesswoman interviewed during the initial trials of the ASLPR was rated at L:0+ despite exceptional facility in comprehending sums of money.

The ASLPR measures general proficiency through the application of global descriptions. This raises at least two crucial issues. First, one has to distinguish the underlying general proficiency from its realization in actual tasks. Thus, a FL learner in a country where door-to-door milkmen do not exist may be unable to write an appropriate note to the milkman to change a daily order but the ability to identify a relationship and the minimum information to be conveyed and to select the appropriate language from his repertoire to carry out the necessary note-writing task should be observed if the learner is to be rated at W:1. The actual task chosen by the interviewer to elicit the underlying proficiency may well differ from learner to learner or situation to situation. Similarly, someone learning English in the Australian outback may not use language entirely appropriately to buy a bus or train ticket but one would expect to see comparable behaviour in some other situation before assigning a rating of S:1-. In other words, it is the underlying proficiency that must be demonstrated and not the fulfilment of an absolutely specified task if a

certain rating is to be assigned. For this reason, the ASLPR is accompanied by recommendations on assessment procedures and not by a standardized set of test materials.

Secondly, the notion of language register and the situational variation of language raises significant issues for a scale claiming to measure general proficiency. On the one hand, if proficiency descriptions are related to particular situations, one can be accused of describing only proficiency in certain registers. On the other hand, language use occurs only in context and "proficiency" unrelated to situations would seem to be a meaningless concept. However, if "general proficiency" means anything, it must mean ability to use the language in the situations regularly encountered in everyday life by people living in the second language environment. Although "everyday situations" are more easily defined by what they are not than by what they are (e.g., not the language of nuclear physics laboratory or a neurology conference), nevertheless they would seem to include those situations necessarily encountered by persons living and meeting their needs in this physical and social world. Thus, "everyday situations" would seem to include, for example, those encountered in moving around one's environment, obtaining food, maintaining health, working or going to school, and so on. Furthermore, no register is entirely discrete or it would not be recognized as part of "the language" and those common elements (e.g., of lexis, syntax, discourse, phonology and functions) that occur across registers and in different situations would seem to be implied in the notion of general or "non-specialist register" (James and Rouve 1973) and to determine the level of the learner's register flexibility, i.e., his ability to operate in new situations involving new register requirements and to modulate his language in response to changes in situation, role, topic, tone or style. It is this concept of general proficiency that underlies the ASLPR and which it seeks to measure.

The parameters of change through the scale reflect this view of general proficiency. In learning a language, the learner (especially if he is a migrant) will first seek to satisfy immediate needs, needs determined either by his need to survive in the L2 environment or by the situations in which the teacher or course designer chooses to place him. As these needs are satisfied, he will reach out to more diverse needs, to establish relationships with others, and to meet vocational needs. Thus, at the lower levels of the scale, the major parameter is the range of situations in which the language can be used. In the middle levels, once sentence formation is more or less complete, the language complexifies through the emergence of modification devices, discourse development, embedding and alternate forms to enable the more precise statement and comprehension of personal meaning and leading, in the upper levels of the scale, to register flexibility. Increasingly through the scale, reference is made to the ability to see beyond surface

meaning and to the ability to cope with increased utterance rates, dialectal variations, and cultural features. Thus, for instance, the ability to perceive humour and irony is distinctive between 4 and 5.

While reference has been made in the definitions to trends in grammatical development, specific items occur only as explanatory examples in the middle column of the scale. Clark (1972:126) deplors the mingling of linguistic and communicative competence criteria in proficiency scales but they have been included in the ASLPR because grammar is a part of language behaviour, language development can be partly described in syntactic terms, and, since one's intuitive judgement of proficiency is influenced by the learner's grammatical accuracy, it seemed preferable to direct the tester's judgement appropriately rather than leave it to chance. Nevertheless, the grammatical examples given do not constitute a checklist since, again, the purpose of the scale is to identify points in a complex stream of development and not to focus predominantly on one component in that stream.

#### 4. APPLYING THE ASLPR

The ASLPR is applied as the rater matches the learner's language behaviour against one of the performance descriptions in the scale. While this approach may be subjective it is not impressionistic and raters are expected to keep the scale in front of them and to carefully match the observed and described behaviours. In order to observe the learner's language proficiency, the rater or, more desirably, an interviewer interviews the learner to elicit a display of his ability to use the language in speaking, listening and reading. For writing, pencil-and-paper exercises must, of course, be done but, as with all four macroskills, the aim is to have the learner demonstrate his maximum language proficiency.

The art of the interviewer (or examiner) is to fully extend the learner without frustrating or flustering him. For the learner, the interview should be informal and relaxed with emphasis falling on his engaging in natural, real-life language activities. As is recommended for the FSI Scale, three broad stages exist in the interview. In the first, the *exploratory* stage, the aim is to put the learner at ease and to provide the interviewer with a general impression of the learner's likely proficiency level. In the second, the *analytical* stage, the interviewer must seek to explore the full extent of the learner's ability, gradually extending him to his linguistic "breaking-point" and exploring, in so doing, his facility in those features that differentiate one ASLPR level from another. In the final, *concluding* stage of the interview, the interviewer rounds off the interview, reverting to activities well within the learner's proficiency level so as to avoid sending him away with a feeling of failure. The time for the whole interview varies depending on the learner's level but, for speaking, listening, reading and the average total time is approximately 20 minutes. The writing test may take longer for more proficient learners but can be administered to a group.

#### 5. USES OF THE ASLPR

The ASLPR aims to measure second language proficiency and not necessarily communicative competence if the latter term is taken in its loosest sense of the ability to communicate. The ability to communicate involves much more than language proficiency and is influenced by such factors as introversion and extroversion, intelligence, education, personal experience, relationship to the other person, and the willingness of the latter to accommodate the speaker's or writer's non-native features. In this sense, it is unlikely that a satisfactory measure of communicative competence will ever be achieved or is, in any case, a proper interest for the language tester. Sollenberger expresses the problem thus:

The person's so-called language proficiency, while it may have been quite accurate in technical skill terms, did not mean effectiveness in communication. In some cases, it may have enabled the person to misrepresent or foul up more effectively. This is to say that you can be a fool in any language... the fact of technical ability to use a foreign language without noticeable accent or grammatical errors (does not) mean that the person has something worth saying. I'm sure we all know people who talk nonsense fluently.

On the other hand I know people who butcher the language, whose accents are atrocious and whose vocabularies are limited. For those reasons we give them a low proficiency rating. Yet, for some reason, some of them are effective communicators.

[Sollenberger 1978:12]

The ASLPR, both potentially and now in practice, has a wide range of uses which, here, can be only summarized:

(i) Most fundamentally, the ASLPR provides a means of expressing a learner's proficiency in a second or foreign language in performance terms that are more directly meaningful to the "consumers" of language tests (e.g., an employer, an educational institution, or the learner himself) than a mark on some more or less arbitrary numerical scale.

(ii) By providing an overall picture of the way in which language proficiency develops from zero to native-like, it provides a conceptual framework within which a language programme or series of language programmes can be coherently planned.

Nevertheless, the ASLPR is not a language syllabus: one needs to differentiate what is taught and repeated at the end of a lesson from what the learner can do out-of-class in real language-use situations. In planning course specifications, it is what is to be presented that is of interest; in defining proficiency, it is ability to use language in real language-use situations. Confusing the two tends to press a false view of language



significant correlations between the ASLPR and the CELT test occur though the extent of these correlations differs and may depend on the types of courses followed and the circumstances under which language learning takes place. Thus, in the studies whose results are summarized in Table 1, quite high correlations were found between the ASLPR and CELT with learners who learned English as a Second Language in functional courses in English-speaking Australia; lower (but still significant) correlations were found with learners following traditional grammar-based courses stressing rote memorization in China where they had little experience in using the language for practical communication purposes (Table 2). The explanation of these differences probably rests in several facts: first, the range of ASLPR ratings across the Chinese students was much narrower than across the Australian learners; second, the ASLPR and CELT are, essentially, testing different, though related, things (*viz.*, on the one hand, practical proficiency and, on the other, knowledge about the language); and third, practical proficiency is influenced by factors other than language learning, factors such as the learner's willingness to "have a go" and his ability to take maximum use of what resources he has for communication purposes. During the French trials of the ASLPR, for example, it was observed (though not quantified) that multilingual High School students in a predominantly "migrant" suburb of Sydney seemed to make more effective use of their limited French repertoire than did their English-mongling classmates. A fourth reason for the differences lies in the greater opportunity the Australian learners had to apply their knowledge in communication activities. Other studies (e.g., Mullen 1978) have shown that indirect and direct tests correlate more closely when used with second language learners who have had frequent opportunity to apply their language knowledge than when used with foreign language learners who develop passive knowledge but have fewer opportunities to apply it in real-life activities. In all, the correlations found between the ASLPR and the CELT test suggest that, where large student numbers are involved, standardized tests could be used for the initial selection of students though too precise interpretation of the results without specific follow-up ratings of individuals on the ASLPR should be avoided.

(vii) The ASLPR can be used for diagnostic purposes. Most simply, it is not the ASLPR that is used diagnostically so much as its associated interview procedures. While the scale is being applied, the rater can use a checklist of relevant course items (e.g., structures, lexis, functions, discourse markers, or phonological features) on which deficiencies identified for subsequent teaching can be marked. The ASLPR is, however, most effectively used for diagnostic purposes as part of a battery of tests. Thus, for example, 220 students at the Guangzhou Institute of Foreign Languages in China were tested in December 1980 using the CELT Structure, Listening and Vocabulary tests, a cloze test, a dictation, and, for a 20% cross-section, the ASLPR (see Ingram 1980). Whereas the CELT test clearly

learning (as rote memory or only "formal" learning) and to over-simplify the teacher's task by reducing it to the practice of the components that make up the ASLPR's proficiency descriptions. For this reason, a clear distinction must be made between *goals* (the ultimate aims definable in terms of the ASLPR) and *objectives* (the teacher's immediate aims in selecting, presenting and practising functions and language exponents and in engaging in other activities). Assessing attainment of objectives entails formative testing in which more traditional, semi-direct and direct instruments can be used; assessing attainment of the proficiency goals entails summative testing for which instruments such as the ASLPR are more suitable.

(iii) In language teaching programmes the ASLPR serves many additional purposes including the statement of goals and attainment, clarifying for learners and others the practical skills courses can develop, providing one basis for streaming, and assessing the suitability of course activities (such as literary study) to particular learners.

(iv) The ASLPR can be used to provide practical measurements of second language skills for many purposes within society. Thus, salary loadings of bilingual employees can be related to real language proficiency; minimum standards for language teacher accreditation can be specified (c.f. Clifford 1978); the language skills of translators and interpreters can be stated; in judicial contexts, a defendant's ability to cope with court proceedings or to understand documents he has signed can be objectively judged; and more accurate judgements about the language adequacy of foreign trained professionals seeking local registration can be made without their language skills being confused with professional criteria. In the last case, for instance, the Queensland Board of Teacher Education has non-native speaking applicants for teacher registration assessed by experts in using the ASLPR; if the applicant is rated at S:4, L:4, R:4, W:4, he can (provided his professional training is adequate) be immediately registered. Below that level, he will be referred to an appropriate ESP or "general purpose" course.

(v) The ASLPR can be used for sociolinguistic research purposes enabling, for instance, a clearer picture to be obtained of a nation's language resources by providing a meaningful way of stating proficiency levels in the population. It has, for example, been used for this purpose in surveys of English language levels and needs in several Australian states.

(vi) The ASLPR, when correlated with available standardized pencil-and-paper tests, can be used to explicate their knowledge-based results in terms of practical proficiency. Thus, if large numbers of students have to be tested, the ASLPR may be too time-consuming to apply but, if correlations between the ASLPR and standardized pencil-and-paper tests do occur, then initial selection of students could be made using the pencil-and-paper tests. In fact, results to date (see Tables 1 and 2 on pp. 56-57) show that

TABLE 1. Pearson Product—Moment ( $r$ ) and Spearman Rank Order ( $\rho$ ) Correlations between ASLPR Macroskill Ratings and the Comprehensive English Language Test (CELT) for Adult Learners of ESL in Australia.

ASLPR	CELT							
	Structure-A			Listening-A		Vocabulary-A		CELT Total
	r	rho		r	rho	r	rho	
Speaking	r/rho							
S		.83	.91	.86	.9	.88	.87	.89 .9
	n	19	19	18	18	19	19	18 18
	s	.000	.001	.000	.001	.000	.001	.000 .001
Listening	r/rho							
L		.80	.9	.86	.9	.85	.83	.87 .88
	n	19	19	18	18	19	19	18 18
	s	.000	.001	.000	.001	.000	.001	.000 .001
Writing	r/rho							
W		.88	.94	.89	.9	.91	.91	.93 .95
	n	19	19	18	18	19	19	18 18
	s	.000	.001	.000	.001	.000	.001	.000 .001
Reading	r/rho							
R		.84	.95	.87	.9	.92	.94	.92 .96
	n	19	19	18	18	19	19	18 18
	s	.000	.001	.000	.001	.000	.001	.000 .001

**r = Pearson Product-Moment Correlation**  
**rho = Spearman Rank Order Correlation**  
**n = number of learners**  
**s = significance level**  
**s of .000 = significance beyond .001 level**

showed the strong structural knowledge of the students and their weak vocabulary, the ASLPR revealed their functional weaknesses reflected in their inability to mobilize their formal knowledge to carry out basic communication tasks appropriately in speaking and writing (e.g., seeking information about the interviewer's country of origin in speaking or noting in writing) or to carry out everyday reading tasks such as getting information on dosage from a medicine bottle label (see Ingram 1980 for a fuller discussion of the diagnostic usefulness of this battery).

## 5. FORMAL TRIALS

## 5.1 Introduction

Since its initial release in February 1979, the ASLPR has been informally assessed in, and feedback obtained from, its everyday use for streaming and assessment purposes throughout the Adult Migrant Education Program

**TABLE 1** Pearson Product-Moment ( $r$ ) and Spearman Rank Order ( $\rho$ ) Correlation between ASLPR Macroskill Ratings and the Comprehensive English Language Test (CELT) for Third Year Students of EFL at the Guangzhou Institute of Foreign Languages, Guangzhou, China.

ASLPR		CELT							
		Structure-A		Listening-A		Vocabulary-A		CELT Total	
		r	rho	r	rho	r	rho	r	rho
Speaking S	r/rho	.57	.59	.50	.52	.3	.26	.51	.55
	n	21	21	21	21	21	21	21	21
	s	.003	.002	.01	.008	.094	.13	.009	.005
Listening L	r/rho	.52	.55	.50	.52	.32	.27	.51	.58
	n	21	21	21	21	21	21	21	21
	s	.008	.005	.01	.008	.081	.121	.01	.003
Writing W	r/rho	.44	.49	.64	.57	.51	.038	.64	.57
	n	21	21	21	21	21	21	21	21
	s	.022	.012	.001	.003	.01	.043	.001	.003
Reading R	r/rho	.32	.43	.41	.45	.14	.01	.31	.3
	n	21	21	21	21	21	21	21	21
	s	.079	.025	.031	.021	.274	.483	.083	.096

**r** = Pearson Product-Moment Correlation  
**rho** = Spearman Rank Order Correlation  
**n** = number of learners  
**s** = significance level

in Australia. Formal trialling has been completed into its validity and reliability when used to rate the proficiency of adult ESL learners and trials are currently being completed into its use with adolescent ESL learners and adolescent and adult FL learners of French, Italian and Japanese.

## 6.2. Validity

Tables 1 and 2 show a satisfactory correlation between the ASLPR and the CELT test, suggesting an acceptable level of immediate pragmatic (or concurrent) validity. The previous discussion on assessing the extent of correlation should be noted: a high or stable degree of correlation should not be expected since tests such as CELT and the ASLPR are essentially testing different though related things, viz., language knowledge, on the one hand, and the mobilization of language knowledge in practical proficiency on the other (cf. Farhady 1979; Mullen 1978). The importance of the use of the ASLPR in a FL learning situation where the learner is likely to have



TABLE 3. KEY HEADINGS FROM  
AUSTRALIAN SECOND LANGUAGE PROFICIENCY RATING (ASLPR)

S:0 Zero Proficiency	Unable to function in the language.	S:0+ Initial Proficiency	Unable to operate only in a very limited capacity within very predictable areas of need.	L:0+ Initial Proficiency	Unable to comprehend the spoken language.	W:0 Zero Proficiency	Unable to function in the written language.	R:0+ Initial Proficiency	Unable to read only a limited range of essential sight words whose forms have been memorized in response to immediate needs.	R:1- Elementary Proficiency	Unable to read short simple sentences and short instructions.	R:1 Minimum Survival Proficiency	Unable to read personal and place names, street signs, office or shop designations, numbers, isolated words and phrases, and short sentences.	R:1+ Survival Proficiency	Unable to read short texts on subjects related to immediate needs.	S:1- Elementary Proficiency	Unable to satisfy immediate needs using learned utterances.	S:1 Minimum Survival Proficiency	Unable to satisfy basic survival needs and minimum courtesy requirements.	L:1- Elementary Proficiency	Unable to comprehend readily only utterances which are thoroughly familiar or are predictable within the areas of immediate survival needs.	L:1 Minimum Survival Proficiency	Unable to comprehend enough to meet basic survival needs.	W:1- Elementary Proficiency	Unable to write with reasonable accuracy short words and brief familiar utterances.	W:1 Minimum Survival Proficiency	Unable to satisfy basic survival needs.	W:1+ Survival Proficiency	Unable to satisfy all survival needs and limited social needs.	S:1+ Survival Proficiency	Unable to satisfy all survival needs and limited social needs.
-------------------------	-------------------------------------	--------------------------	--	--------------------------	---	----------------------	---	--------------------------	--	-----------------------------	---	----------------------------------	---	---------------------------	--	-----------------------------	---	----------------------------------	---	-----------------------------	---	----------------------------------	---	-----------------------------	---	----------------------------------	---	---------------------------	--	---------------------------	--

S:2 Minimum Social Proficiency	Abble to satisfy routine social demands and limited work requirements.	S:3 Minimum Vocational Proficiency	Abble to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social and vocational topics.	L:2 Minimum Social Proficiency	Abble to understand in routine social situations and limited work situations.	L:3 Minimum Vocational Proficiency	Abble to comprehend sufficient proficiency	W:2 Minimum Social Proficiency	Abble to satisfy routine social demands and limited work requirements.	W:3 Minimum Vocational Proficiency	Abble to write with sufficient accuracy in structures and spelling to meet all social needs and basic work needs.	R:2 Minimum Social Proficiency	Abble to read simple prose, in a form equivalent to typescript or printing, on subjects within a familiar context.	R:3 Minimum Vocational Proficiency	Abble to read standard news-paper items addressed to the general reader, routine correspondence, reports and technical material in his special field, and other everyday materials (e.g. best-selling novels and similar recreational literature).	L:4 Vocational Proficiency	Can comprehend easily and accurately in all personal and social contexts and in all texts relevant to own experience.	W:4 Vocational Proficiency	Abble to write fluently and accurately on all levels normally pertinent to personal, social, academic or vocational needs.	R:4 Vocational Proficiency	Abble to read all styles and forms of the language pertinent to personal, vocational, social, academic or vocational needs.	L:5 Native-like Proficiency	Listening proficiency equivalent to that of a native speaker of the same socio-cultural variety.	W:5 Native-like Proficiency	Written proficiency equivalent to that of a native speaker of the same socio-cultural variety.	R:5 Native-like Proficiency	Reading proficiency equivalent to that of a native speaker of the same socio-cultural variety.
--------------------------------	--	------------------------------------	---	--------------------------------	---	------------------------------------	--	--------------------------------	--	------------------------------------	---	--------------------------------	--	------------------------------------	--	----------------------------	---	----------------------------	--	----------------------------	---	-----------------------------	--	-----------------------------	--	-----------------------------	--

NOTE: This table provides only the key headings used in the ASLPR. A copy of the full scale can be obtained by writing to Dr. D.B. Ingram, Darwin Community College, School of General Studies, Winnellie, N.T. Australia.

is experience of using the language for practical communicative purposes indicated by the lower correlations between the ASLPR and CELF in the China Study (see Tables 1 and 2).

Construct and content validity have been ensured by the procedures which the ASLPR has been developed, especially by the continual examination of its proficiency descriptions by comparing them with the actual language behaviour of some hundreds of learners. Face validity seems to be demonstrated by the rapid adoption of the ASLPR throughout the Adult Migrant Education Program in Australia, in foreign language programmes in tertiary institutions, by professional registration authorities such as the Queensland Board of Teacher Education, in sociolinguistic research, and, not least, in courts of law.

### 3 Reliability

examining reliability, the formal trials were interested in the extent to which raters could reliably interpret and apply the proficiency descriptions assessing learners. Specifically, the trials examined:

- (i) inter-rater reliability (i.e., the extent to which 21 Australian raters, including 20 native English speakers and one non-native, and 15 non-native speaking Chinese raters could rate the same 16 learners in the same way as the developers of the scale);
- ii) Intra-rater reliability (i.e., the extent to which the 21 Australian raters could rate the same learners alike on two different occasions).

The method adopted to assess reliability involved the videotaping of ten learners distributed through the proficiency scale from zero to five-like. Speaking, listening and reading skills were shown on the half-hour videos while written scripts were obtained from the learners to illustrate their writing skills. The videos and scripts were then shown in random order to 20 experienced ESL teachers in Sydney, Melbourne, and Adelaide, first, in April-May 1980 and again in March-April 1981. On the occasion, care was taken to ensure that the teachers did not take out the room with them any record of the ratings they had assigned and, subsequently, the twelve months lapse before the re-run of the trials ensured that memory of the earlier ratings assigned did not influence the second rating. In January 1981, the same videos and scripts in the same random order were also shown to 15 experienced Chinese teachers of English from universities around China who were studying applied linguistics at the Guangzhou Institute of Foreign Languages. The Chinese teachers were all non-native speakers of English with proficiencies in all crosskills around 3 on the ASLPR. Before seeing the trial material, they had received a ten-hour introduction to the use of the ASLPR as a part of their course. This procedure has been repeated with two groups

of native speaking teachers in the United States but these results have yet to be analyzed.

The results of the reliability trials are reported in full in Ingram 1982. Here only a brief summary will be presented though readers may obtain further details from the author.

Correlations between each teacher's and the researchers' ratings were assessed using the Spearman Rank Order Correlation ( $\rho$ ) and the Pearson Product-Movement Correlation ( $r$ ). For the Australian teachers, correlation coefficients range as follows:

*Speaking* :  $\rho$  ranges .91 to .99 with a mean of .96;  $r$  ranges .94 to .99 with a mean of .96  
*Listening* :  $\rho$  ranges .89 to .98 with a mean of .94;  $r$  ranges .9 to .99 with a mean of .94  
*Writing* :  $\rho$  ranges .23 to .997 with a mean of .96;  $r$  ranges .93 to .99 with a mean of .96  
*Reading* :  $\rho$  ranges .91 to .99 with a mean of .96;  $r$  ranges .91 to .98 with a mean of .95.

For the 15 Chinese teachers, the correlations were only slightly lower:

*Speaking* :  $\rho$  ranges .88 to .98 with a mean of .95;  $r$  ranges .83 to .98 with a mean of .93  
*Listening* :  $\rho$  ranges .88 to .96 with a mean of .93;  $r$  ranges .89 to .97 with a mean of .94  
*Writing* :  $\rho$  ranges .88 to .98 with a mean of .95;  $r$  ranges .87 to .98 with a mean of .94  
*Reading* :  $\rho$  ranges .88 to .98 with a mean of .95;  $r$  ranges .83 to .98 with a mean of .93.

Clearly, these figures suggest high inter-rater reliability. In other words, in this study, both native and non-native speaking teachers were able to interpret and apply the ASLPR in a very similar manner to the developers of the scale.

Furthermore, intra-rater reliability coefficients were equally high, 17 teachers rating the same learners in essentially the same way on two separate occasions. Specifically, correlation coefficients between the ratings assigned on two separate occasions twelve months apart were:

*Speaking* :  $\rho$  ranges .93 to .99 with a mean of .97;  $r$  ranges .93 to .99 with a mean of .97  
*Listening* :  $\rho$  ranges .91 to .98 with a mean of .96;  $r$  ranges .92 to .98 with a mean of .96  
*Writing* :  $\rho$  ranges .92 to .98 with a mean of .96;  $r$  ranges .91 to .99 with a mean of .96  
*Reading* :  $\rho$  ranges .87 to .98 with a mean of .96;  $r$  ranges .91 to .98 with a mean of .96.



To sum up, formal trials have shown that the ASLPR can claim acceptable levels of validity while high levels of inter- and intra-rater reliability have been demonstrated. In other words, the ASLPR does seem to make valid statements about a learner's language proficiency in a way acceptable to users of such test results and, equally importantly, raters are able to interpret the proficiency descriptions that constitute the scale and seem able to apply them reliably in assessing second language proficiency.

#### CONCLUSION

A rapid acceptance of the ASLPR in Australia suggests that, at least in that country, it is filling a long-felt need for a practical way of measuring second language proficiency of bilinguals. While second and foreign language teaching focuses on the development of practical skills, there is a need for an instrument that enables goals to be specified in terms of proficiency and for the learner's skills to be described in terms of what can actually do in the language. As the results of the formal trials of scale in China have shown, its usefulness is by no means limited to Australia or to native speaking raters. Further study is to be made of review techniques for use by non-native speakers and the results of formal trials with foreign language learners of languages other than English have yet to be finalized, though initial results suggest reliability coefficients similar to those outlined above. In all, it seems safe to conclude that the ASLPR has a high level of validity and reliability and that it can confidently be used wherever a meaningful statement of a learner's practical skills in the second language is required.

#### REFERENCES

- Geis, P.J. 1979. TOEFL in recent years. In: Spolsky (ed). pp. 101-103.
- Yoffe, W.H. and A.T. Sharon. 1970. A comparison of scores earned on the test of English as a foreign language by native American college students and foreign applicants to U.S. colleges. [Princeton, N.J.: Educational Testing Service; 1970] (mimeograph).
- Clark, J.L.D. 1972. *Foreign language testing: theory and practice*. Philadelphia, Pa.: Center for Curriculum Development.
- , 1978. Direct vs. semi-direct tests of speaking ability. Paper to the conference on *The direct testing of speaking proficiency: Theory and applications*, Georgetown University, Washington, March 14th-15th, 1978. (mimeograph).
- Clark, R.T. 1978. Reliability and validity of language aspects contributing to oral proficiency of prospective teachers of German. Paper to

the conference on *The Direct Testing of Speaking Proficiency: Theory and Application*, Georgetown University, Washington, March 14th-15th. (mimeograph).

Farhady, H. 1979. The disjunctive fallacy between discrete-point and integrative tests. *TESOL Quarterly*, 13.3, September, 1979. pp. 347-357.

Foreign Service Institute School of Language Studies 1968. "Absolute language proficiency ratings". Washington, U.S. Department of State 1968 (mimeograph), rpt. Clark 1972:122-123.

Graham, S.L. 1978. Using the FSI interview as a diagnostic evaluation instrument. Paper to the conference on *The Direct Testing of Speaking Proficiency: Theory and Application*, Georgetown University, Washington, March 14th-15th, 1978. (mimeograph).

Harris, D.P. and L.A. Palmer. 1970. *A comprehensive English language test for Speakers of English as a Second Language*. New York: McGraw Hill.

Ingram, D.E. 1977. Language proficiency at Matriculation. Paper to the 1977 Congress of the Applied Linguistics Association of Australia, University of Melbourne, August 29th-31st, 1977, reprinted in *Australian Review of Applied Linguistics*, 1.1 20-32.

—, 1978. An Applied Linguistic Study of Advanced Language Learning. Thesis for the degree of Doctor of Philosophy, University of Essex, April 1978.

—, 1979. "Methodology" in the Teachers' Manual of the "On-Arrival" segment of the Adult Migrant Education Program Canberra. Department of Immigration and Ethnic Affairs; 1979, Chapter A.

—, 1980. Report on the proficiency assessment of second and third year students at the Guangzhou Institute of Foreign Languages, Guangzhou, China, November-December 1980 (mimeograph).

—, 1982. Report on the Formal Trialling of the Australian Second Language Proficiency Ratings. Report of a research project funded by the Migrant Education Branch of the Australian Department of Immigration and Ethnic Affairs, Canberra, 1982 (mimeograph).

James, C.V. and S. Rouve, 1973. *Survey of curricula and performance in modern languages*. London: CILT.

Lado, R. 1978. Scope and limitations of interview-based language testing: Are we asking too much of the interview? Paper to the conference on *The Direct Testing of Speaking Proficiency: Theory and Application*, Georgetown University, Washington, March 14th-15th, (mimeograph).

Lovelace, W. 1978. Interview testing in non-European languages. Paper to the conference on *The Direct Testing of Speaking Proficiency: Theory and Application*, Georgetown University, Washington, March 14th-15th, (mimeograph).

Mullen, K.A. 1978. Determining the effect of uncontrolled sources of error in a direct test of oral proficiency and the capability of the procedure to detect improvement following classroom instruction. Paper to the conference on *The Direct Testing of Speaking Proficiency: Theory and Application*, Georgetown University, Washington, March 14th-15th, 1978. (mimeograph).

Oller, J.W. and B. Spolsky. 1979. The test of English as a foreign language. In: Spolsky (ed) 1979, pp. 92-100.

Sollenberger, H.E. 1978. Development and current use of the FSI Oral Interview Test. Paper to the conference on *The Direct Testing of Speaking Proficiency: Theory and Application*, Georgetown University Washington, March 14th-15th, 1978. (mimeograph).

Spolsky, B. 1979, *Papers in applied linguistics: advances in language testing*. Series 1. Arlington, Virginia; Center for Applied Linguistics.

Wilds, C.P. 1965: The measurement of speaking and reading proficiency in a foreign language. [mimeograph supplied by Foreign Service Institute School of Language Studies; c. 1965].

## Bilingualism through schooling in India

E. ANNAMALAI  
Central Institute of Indian Languages  
Mysore, India

**ABSTRACT.** This is an investigative study into the Indian school bilingualism. The effects of the application of the famous three language formula (TLF) are discussed. School bilingualism for the majority group is not in any modern Indian language besides Hindi (and English). In such a case the role of schools in spreading bilingualism is limited (one case is of English, which is learnt in formal setting of education only). It is proposed that national bilingualism can only be achieved if regional languages assume importance outside their states, and secondly, if the present state of teaching methods of SL and FL teaching are improved by bringing radical changes in the ideology and methods of language teaching in schools.

Bilingualism<sup>1</sup> in India is characterized by its social acceptance as a normal societal phenomenon. To be a bilingual is not considered to be a deviant or exceptional behaviour, though the percentage of bilinguals in the country with about 200 languages is hardly 10% (9.7% according to 1961 Census)<sup>2</sup> of the total population. People learn another language as part of their socialization process whenever there is a need for it for communicative, economic, cultural or religious purposes and they maintain the languages for use in different domains without functional conflict, though with formal convergence of their grammars as a result of constant switching between them.

The percentage of bilingualism is considerable when compared with the percentage of literacy in India, which was 24.02% in 1961. For about every two literate persons, there is one bilingual. It does not, however, mean that bilinguals are a subset of literate persons. Illiteracy in India does not come in the way of one becoming a bilingual. It is likely that